

Behaviors of φ -exponential distributions in Wasserstein geometry and an evolution equation

Asuka Takatsu *

Abstract

A φ -exponential distribution is a generalization of an exponential distribution associated to functions φ in an appropriate class, and the space of φ -exponential distributions has a dually flat structure. We study features of the space of φ -exponential distributions, such as the convexity in Wasserstein geometry and the stability under an evolution equation. From this study, we provide the new characterizations to the space of Gaussian measures and the space of q -Gaussian measures.

1 Introduction

A Gaussian measure is an exponential distribution on \mathbb{R}^d with mean and covariance matrix parameters. The space of Gaussian measures has good behaviors such as the flatness in a dual structure which is a Riemannian metric with a pair of connections being orthogonal to each other, the convexity in Wasserstein geometry which is a metric geometry on the space of probability measures, and the stability under the linear evolution equation called the Fokker–Planck equation. As for the validity of these behaviors, for instance, the dually flat structure on the space of Gaussian measures leads the Cramér–Rao lower bound for any unbiased estimator.

Recently, Naudts [4] introduced the space of probability measures, which is a generalization of the space of Gaussian measures equipped with a dually flat structure, and generalized the Cramér–Rao lower bound, where each probability measure is determined by the inverse function of the function $t \mapsto \int_1^t 1/\varphi(s)ds$ on $(0, \infty)$ for an increasing, positive, continuous function φ on $(0, \infty)$. The probability measure is called a *φ -exponential distribution* and coincides with a usual exponential distribution when $\varphi(s) = s$. The φ -exponential distribution on \mathbb{R}^d has a finite second moment if φ has a certain limiting behavior at 0 and ∞ , which corresponds to the instances of functions φ given by $\varphi(s) = s^q$

*Graduate School of Mathematics, Nagoya University, Nagoya 464-8602, Japan & Institut des Hautes Études Scientifiques, Bures-sur-Yvette 91440, France. (takatsu@math.nagoya-u.ac.jp)

Mathematics Subject Classification (2010): 60D05, 94A17.

keywords: φ -exponential distribution, q -Gaussian measure, Wasserstein geometry, evolution equation.

with $q \in (0, (d+4)/(d+2))$. When $\varphi(s) = s^q$ except for $q = 1$, the φ -exponential distribution with mean and covariance matrix parameters is called a q -Gaussian measure, which is a power-law distribution.

The space of q -Gaussian measures behaves well due to the convenience of power-law distributions. Indeed, the space of q -Gaussian measures is convex in Wasserstein geometry and stable under the nonlinear evolution equation of porous medium type (see [5, Proposition 5] and [7, Theorem A]). We prove that the convexity and the stability are features only of the space of Gaussian measures and the space of q -Gaussian measures. In other words, the condition $\varphi(s) = s^q$ with $q \in (0, (d+4)/(d+2))$ is a necessary and sufficient condition for the space of φ -exponential distributions on \mathbb{R}^d with mean and covariance matrix parameters to have each of the two features.

The paper is organized as follows. In the next section, we give definitions and properties about Wasserstein geometry and the φ -exponential function. Section 3 concerns a condition for φ -exponential distributions to have a finite second moment (Proposition 3.2). The two spaces of φ -exponential distributions with mean and covariance matrix parameters are introduced in Section 4, one of which has a dually flat structure and the other is convex in Wasserstein geometry (Proposition 4.3). We discuss when the two spaces coincide with each other (Theorem 4.1). In Section 5, we consider the stability of the two spaces under a certain evolution equation (Theorem 5.1 and Corollary 5.2).

Acknowledgements. The author would like to thank Shin-ichi Ohta for many valuable advice and discussions. She is also grateful to Sumio Yamada for helpful comments. She is partially supported by JSPS-IHÉS (EPDI) fellowship.

2 Preliminaries

2.1 Wasserstein geometry

Let us briefly recall properties of (L^2) -Wasserstein geometry over \mathbb{R}^d . See [8, 9] and references therein for further information.

In this paper, any measure is always supposed to be a Borel measure. A measure μ on \mathbb{R}^d is said to be *absolutely continuous* with respect to the Lebesgue measure \mathcal{L}^d on \mathbb{R}^d if there exists a measurable function ρ on \mathbb{R}^d such that $\mu = \rho \mathcal{L}^d$. For a measurable map F on \mathbb{R}^d and a measure μ on \mathbb{R}^d , the *push-forward measure* $F_\# \mu$ of μ thorough F is a measure on \mathbb{R}^d defined by $F_\# \mu[B] := \mu[F^{-1}(B)]$ for all Borel sets $B \subset \mathbb{R}^d$. Given two measures $\mu = \rho \mathcal{L}^d, \nu = \sigma \mathcal{L}^d$ and a differentiable map F on \mathbb{R}^d , $\nu = F_\# \mu$ is equivalent to $\rho = \sigma(F) \det(dF)$ μ -almost everywhere, where dF is the total differential of F . For any probability measures μ and ν on \mathbb{R}^d , a *coupling* π of μ and ν is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν . In other words, we have $\pi[B \times \mathbb{R}^d] = \mu[B]$ and $\pi[\mathbb{R}^d \times B] = \nu[B]$ for all Borel sets $B \subset \mathbb{R}^d$. We denote by \mathcal{P}_2 the set of probability measures μ on \mathbb{R}^d with finite second moments, namely

$$\int_{\mathbb{R}^d} |x|^2 d\mu(x) < \infty.$$

Definition 2.1 We define the *Wasserstein distance* between μ and ν in \mathcal{P}_2 by

$$W_2(\mu, \nu) := \inf \left\{ \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\pi(x, y) \right)^{\frac{1}{2}} \mid \pi : \text{a coupling of } \mu \text{ and } \nu \right\}. \quad (2.1)$$

The function W_2 is indeed a distance function on \mathcal{P}_2 . The metric space (\mathcal{P}_2, W_2) is called the *Wasserstein space*, where any two points are joined by a geodesic. Here a *geodesic* is a distance-minimizing curve and has a constant speed. It is known that any Wasserstein geodesic does not branch, that is, if two geodesics have a common interval, then they are subintervals of one geodesic. Moreover, a coupling achieving the infimum of (2.1) always exists, which is called an *optimal coupling*. For these facts, see [9, Chapters 6,7]. We refer to the results of Knott–Smith [3] and Brenier [2] about optimal couplings. We denote by id the identity map on \mathbb{R}^d .

Theorem 2.2 ([8, Theorems 2.12, 2.16]) *Let $\mu, \nu \in \mathcal{P}_2$ and μ be absolutely continuous with respect to the Lebesgue measure. If a proper lower semi-continuous convex function ϕ on \mathbb{R}^d satisfies $[\nabla \phi]_{\#} \mu = \nu$, then $[\text{id} \times \nabla \phi]_{\#} \mu$ is a unique optimal coupling of μ and ν , and $\{[(1-t)\text{id} + t\nabla \phi]_{\#} \mu\}_{t \in [0,1]}$ is a unique Wasserstein geodesic from μ to ν .*

2.2 φ -exponential function

We summarize definitions and properties of the φ -exponential function. For further details, we refer to [4] and references therein.

For an increasing, positive, continuous function φ on $(0, \infty)$, the φ -logarithmic function is defined by

$$\ln_{\varphi}(t) := \int_1^t \frac{1}{\varphi(s)} ds,$$

which is increasing, concave and C^1 on $(0, \infty)$. The constants l_{φ} and L_{φ} are respectively defined as the infimum and the supremum of \ln_{φ} , that is,

$$l_{\varphi} := \inf_{t>0} \ln_{\varphi}(t) = \lim_{t \downarrow 0} \ln_{\varphi}(t) \in [-\infty, 0), \quad L_{\varphi} := \sup_{t>0} \ln_{\varphi}(t) = \lim_{t \uparrow \infty} \ln_{\varphi}(t) \in (0, +\infty].$$

The function \ln_{φ} has the inverse function, which is called the φ -exponential function and is defined on $(l_{\varphi}, L_{\varphi})$. This inverse function can be extended to all of \mathbb{R} as

$$\exp_{\varphi}(\tau) := \begin{cases} 0 & \text{for } \tau \leq l_{\varphi}, \\ \ln_{\varphi}^{-1}(\tau) & \text{for } \tau \in (l_{\varphi}, L_{\varphi}), \\ \infty & \text{for } \tau \geq L_{\varphi}, \end{cases}$$

which is C^1 on $(l_{\varphi}, L_{\varphi})$ and its derivative is given by

$$\frac{d}{d\tau} \exp_{\varphi}(\tau) = \varphi(\exp_{\varphi}(\tau)).$$

We mention that if $\varphi(0) = 0$ then the function \exp_{φ} is always C^1 on $(-\infty, L_{\varphi})$.

The case of $\varphi(s) = s$ is the most fundamental case related to the Boltzmann–Gibbs statistics and the Fokker–Planck equation, where the φ -exponential function coincides with the usual exponential function. Another important case is that $\varphi(s) = s^q$, where the φ -exponential function is the power function given by

$$\exp_q(\tau) := [1 + (1 - q)\tau]_+^{\frac{1}{1-q}},$$

where we set $[\tau]_+ := \max\{\tau, 0\}$ for $\tau \in \mathbb{R}$ and $0^a := \infty$ for $a < 0$. This case is related to the Tsallis statistics and the nonlinear evolution equation of porous medium type.

We introduce a class of increasing, positive, continuous functions on $(0, \infty)$.

Definition 2.3 For any $a \in \mathbb{R}$, we define $\mathcal{O}(a)$ to be the set of all increasing, positive, continuous functions φ on $(0, \infty)$ such that $\max\{\delta_\varphi, \delta^\varphi\} < a$, where we set

$$\delta_\varphi := \inf \left\{ \delta \in \mathbb{R} \mid \lim_{s \downarrow 0} \frac{s^{1+\delta}}{\varphi(s)} \text{ exists} \right\}, \quad \delta^\varphi := \inf \left\{ \delta \in \mathbb{R} \mid \lim_{s \uparrow \infty} \frac{s^{1+\delta}}{\varphi(s)} = \infty \right\}.$$

It is trivial that $\mathcal{O}(a) \subset \mathcal{O}(b)$ if $a < b$. We define the constant δ'_φ , expressing the order of \ln_φ at 0, by

$$\delta'_\varphi := \inf \left\{ \delta \in \mathbb{R} \mid \lim_{t \downarrow 0} t^\delta \ln_\varphi(t) \text{ exists} \right\}.$$

We refer to the relation between δ_φ and δ'_φ .

Lemma 2.4 For any $\varphi \in \mathcal{O}(a)$ with some $a \in \mathbb{R}$, we have $\delta'_\varphi \leq \max\{\delta_\varphi, 0\}$.

Proof. Given any $\varepsilon > 0$, there exists $C > 0$ such that

$$0 \leq \frac{s^{1+\max\{\delta_\varphi, 0\}+\varepsilon}}{\varphi(s)} \leq C$$

holds for any $s \in (0, 1)$. Set $\delta := \max\{\delta_\varphi, 0\} + 2\varepsilon$. Then for $t \in (0, 1)$, we have

$$0 \leq -t^\delta \ln_\varphi(t) = t^\delta \int_t^1 \frac{1}{\varphi(s)} ds \leq \int_t^1 \frac{s^\delta}{\varphi(s)} ds \leq C \int_t^1 s^{-1+\varepsilon} ds = \frac{C}{\varepsilon} (1 - t^\varepsilon) < \frac{C}{\varepsilon} < +\infty,$$

proving $\delta'_\varphi \leq \max\{\delta_\varphi, 0\} + 2\varepsilon$. Since $\varepsilon > 0$ is arbitrary, we have $\delta'_\varphi \leq \max\{\delta_\varphi, 0\}$. \square

3 Condition of φ

This section is devoted to the study of φ -exponential distributions with mean and covariance matrix parameters. We first observe behaviors of the following function f_φ .

Lemma 3.1 For any $\varphi \in \mathcal{O}(1)$, we set the function and the constant as

$$f_\varphi(p, \lambda) := \int_0^{\exp_\varphi(\lambda)} \left| (\lambda - \ln_\varphi(t))^p \frac{t}{\varphi(t)} \right| dt = \int_0^{\exp_\varphi(\lambda)} (\lambda - \ln_\varphi(t))^p \frac{t}{\varphi(t)} dt,$$

$$p_\varphi := \begin{cases} \frac{1}{\max\{\delta_\varphi, \delta^\varphi\}} - 1 & \text{if } \max\{\delta_\varphi, \delta^\varphi\} > 0, \\ \infty & \text{otherwise.} \end{cases}$$

- (1) For any $(p, \lambda) \in (-1, p_\varphi) \times (l_\varphi, L_\varphi)$, $f_\varphi(p, \lambda)$ is well-defined.
- (2) The function $\lambda \mapsto f_\varphi(p, \lambda)$ on (l_φ, L_φ) is C^0 for $p \in (-1, 0)$ and C^1 for $p \in [0, p_\varphi)$.
- (3) The function $p \mapsto f_\varphi(p, \lambda)$ on $(-1, p_\varphi)$ is C^∞ for $\lambda \in (l_\varphi, L_\varphi)$.
- (4) For any $p \in (-1, p_\varphi)$, $\lim_{\lambda \downarrow l_\varphi} f_\varphi(p, \lambda) = 0$ and $\lim_{\lambda \uparrow L_\varphi} f_\varphi(p, \lambda) = \infty$ hold.

Proof. Note that for any $p \in (-1, p_\varphi)$, we have $1/(p+1) > \max\{\delta_\varphi, \delta^\varphi, \delta'_\varphi\}$.

(1): Fix any $\lambda \in (l_\varphi, L_\varphi)$ and take $\varepsilon \in (0, \lambda - l_\varphi)$. Given any $p \in (-1, 0)$, we deform

$$\begin{aligned} f_\varphi(p, \lambda) &= \int_0^{\exp_\varphi(\lambda-\varepsilon)} (\lambda - \ln_\varphi(t))^p \frac{t}{\varphi(t)} dt + \int_{\exp_\varphi(\lambda-\varepsilon)}^{\exp_\varphi(\lambda)} (\lambda - \ln_\varphi(t))^p \frac{t}{\varphi(t)} dt \\ &= \int_0^{\exp_\varphi(\lambda-\varepsilon)} (\lambda - \ln_\varphi(t))^p \frac{t}{\varphi(t)} dt + \int_0^\varepsilon s^p \exp_\varphi(\lambda - s) ds \\ &\leq \varepsilon^p \int_0^{\exp_\varphi(\lambda-\varepsilon)} \frac{t}{\varphi(t)} dt + \exp_\varphi(\lambda) \int_0^\varepsilon s^p ds, \end{aligned}$$

where we use the change of variables formula for $s = \lambda - \ln_\varphi(t)$ in the second line, and the monotonicity of \ln_φ and \exp_φ in the inequality. The integrability of the last line follows from the conditions that $\varphi \in \mathcal{O}(1)$ and $p > -1$.

For any $p \in [0, p_\varphi)$, we fix δ as

$$\max\{\delta_\varphi, \delta^\varphi, \delta'_\varphi\} < \delta := \frac{1}{2} \left(\max\{\delta_\varphi, \delta^\varphi, \delta'_\varphi\} + \frac{1}{p+1} \right) < \frac{1}{p+1}$$

and deform $f_\varphi(p, \lambda)$ as

$$f_\varphi(p, \lambda) = \int_0^{\exp_\varphi(\lambda)} (t^\delta \lambda - t^\delta \ln_\varphi(t))^p \frac{t^{1+\delta}}{\varphi(t)} \cdot t^{-\delta(p+1)} dt.$$

Since $\delta(p+1) < 1$ holds and the function $t \mapsto (t^\delta \lambda - t^\delta \ln_\varphi(t))^p t^{1+\delta}/\varphi(t)$ is bounded on $(0, \exp_\varphi(\lambda))$, $f_\varphi(p, \lambda)$ is well-defined.

(2), (3): The assertions follow from Lebesgue's dominated convergence theorem.

(4): The assertion $\lim_{\lambda \downarrow l_\varphi} f_\varphi(p, \lambda) = 0$ also follows from Lebesgue's dominated convergence theorem. Integrating by part with the condition $1/(p+1) > \delta'_\varphi$ yields that

$$\begin{aligned} f_\varphi(p, \lambda) &= \frac{-t}{p+1} (\lambda - \ln_\varphi(t))^{p+1} \Big|_0^{\exp_\varphi(\lambda)} + \frac{1}{p+1} \int_0^{\exp_\varphi(\lambda)} (\lambda - \ln_\varphi(t))^{p+1} dt \\ &\geq 0 + \frac{1}{p+1} \int_0^{\exp_\varphi(\lambda)} \left(\frac{\exp_\varphi(\lambda) - t}{\varphi(\exp_\varphi(\lambda))} \right)^{p+1} dt \\ &= \frac{\exp_\varphi(\lambda)}{(p+1)(p+2)} \left(\frac{\exp_\varphi(\lambda)}{\varphi(\exp_\varphi(\lambda))} \right)^{p+1}, \end{aligned}$$

where the inequality follows from the concavity of \ln_φ , that is,

$$\ln_\varphi(t) \leq \ln_\varphi(\exp_\varphi(\lambda)) + \frac{t - \exp_\varphi(\lambda)}{\varphi(\exp_\varphi(\lambda))} = \lambda + \frac{t - \exp_\varphi(\lambda)}{\varphi(\exp_\varphi(\lambda))}.$$

The condition $1/(p+1) > \delta^\varphi$ leads that

$$\lim_{\lambda \uparrow L_\varphi} \exp_\varphi(\lambda) \left(\frac{\exp_\varphi(\lambda)}{\varphi(\exp_\varphi(\lambda))} \right)^{p+1} = \lim_{s \uparrow \infty} \left(\frac{s^{1+\frac{1}{p+1}}}{\varphi(s)} \right)^{p+1} = \infty,$$

proving $\lim_{\lambda \uparrow L_\varphi} f_\varphi(p, \lambda) = \infty$. \square

Let $S(d, \mathbb{R})_+$ be the set of symmetric positive definite matrices of size d , and $\{e_i\}_{i=1}^d$ be the standard basis on \mathbb{R}^d . Given any $V \in S(d, \mathbb{R})_+$, let $V^{1/2}$ be the symmetric positive definite matrix such that $V^{1/2} \cdot V^{1/2} = V$ and $|x|_V^2 := \langle x, V^{-1}x \rangle$ for $x \in \mathbb{R}^d$.

Proposition 3.2 *For any $\varphi \in \mathcal{O}(2/(d+2))$ with $d \geq 2$, there exist constants $\lambda \in (l_\varphi, L_\varphi)$ and $c > 0$ such that*

$$\int_{\mathbb{R}^d} \exp_\varphi(\lambda - c|x - v|_V^2) d\mathcal{L}^d = 1, \quad (3.1)$$

$$\int_{\mathbb{R}^d} \langle x, e_i \rangle \exp_\varphi(\lambda - c|x - v|_V^2) d\mathcal{L}^d = \langle v, e_j \rangle, \quad (3.2)$$

$$\int_{\mathbb{R}^d} \langle x - v, e_i \rangle \langle x - v, e_j \rangle \exp_\varphi(\lambda - c|x - v|_V^2) d\mathcal{L}^d = \langle e_i, V e_j \rangle. \quad (3.3)$$

Proof. Note that the condition $\varphi \in \mathcal{O}(2/(d+2))$ with $d \geq 2$ implies that $1 \leq d/2 < p_\varphi$. Let $\Gamma(\cdot)$ be the Gamma function. Applying the change of variables formula first for $y = c^{1/2}V^{-1/2}(x - v)$ then from Cartesian to polar coordinates and for $t = \exp_\varphi(\lambda - |y|^2)$, we have

$$\begin{aligned} \int_{\mathbb{R}^d} \exp_\varphi(\lambda - c|x - v|_V^2) d\mathcal{L}^d &= f_\varphi\left(\frac{d-2}{2}, \lambda\right) \cdot \left(\frac{\pi}{c}\right)^{\frac{d}{2}} \frac{\sqrt{\det V}}{\Gamma(d/2)}, \\ \int_{\mathbb{R}^d} \langle x, e_i \rangle \exp_\varphi(\lambda - c|x - v|_V^2) d\mathcal{L}^d &= f_\varphi\left(\frac{d-2}{2}, \lambda\right) \cdot \left(\frac{\pi}{c}\right)^{\frac{d}{2}} \frac{\sqrt{\det V}}{\Gamma(d/2)} \cdot \langle v, e_i \rangle, \\ \int_{\mathbb{R}^d} \langle x - v, e_i \rangle \langle x - v, e_j \rangle \exp_\varphi(\lambda - c|x - v|_V^2) d\mathcal{L}^d &= f_\varphi\left(\frac{d}{2}, \lambda\right) \cdot \left(\frac{\pi}{c}\right)^{\frac{d}{2}} \frac{\sqrt{\det V}}{\Gamma(d/2)} \cdot \frac{\langle e_i, V e_j \rangle}{cd}, \end{aligned}$$

which are well-defined by Lemma 3.1. This means that the pair (λ, c) satisfying

$$\frac{f_\varphi((d-2)/2, \lambda)^{\frac{d+2}{2}}}{f_\varphi(d/2, \lambda)^{\frac{d}{2}}} = (d\pi)^{-\frac{d}{2}} \frac{\Gamma(d/2)}{\sqrt{\det V}}, \quad c = \frac{f_\varphi(d/2, \lambda)}{df_\varphi((d-2)/2, \lambda)} \quad (3.4)$$

is the desired one. If the continuous function $F_\varphi(p, \lambda) := f_\varphi(p-1, \lambda)^{p+1}/f_\varphi(p, \lambda)^p$ on $(0, p_\varphi) \times (l_\varphi, L_\varphi)$ satisfies

$$\lim_{\lambda \downarrow l_\varphi} F_\varphi\left(\frac{d}{2}, \lambda\right) = 0, \quad \lim_{\lambda \uparrow L_\varphi} F_\varphi\left(\frac{d}{2}, \lambda\right) = \infty, \quad (3.5)$$

then the intermediate value theorem guarantees the existence of such a pair (λ, c) .

The rest is to prove (3.5). For any $(p, \lambda) \in (-1, p_\varphi) \times (l_\varphi, L_\varphi)$, Hölder's inequality yields that

$$\begin{aligned} \frac{\partial}{\partial p} f_\varphi(p, \lambda) &= \int_0^{\exp_\varphi(\lambda)} \ln(\lambda - \ln_\varphi(t)) \frac{(\lambda - \ln_\varphi(t))^p t}{\varphi(t)} dt \\ &\leq \left(\int_0^{\exp_\varphi(\lambda)} \ln^2(\lambda - \ln_\varphi(t)) \frac{(\lambda - \ln_\varphi(t))^p t}{\varphi(t)} dt \right)^{\frac{1}{2}} \left(\int_0^{\exp_\varphi(\lambda)} \frac{(\lambda - \ln_\varphi(t))^p t}{\varphi(t)} dt \right)^{\frac{1}{2}} \\ &= \left(\frac{\partial^2}{\partial p^2} f_\varphi(p, \lambda) \right)^{\frac{1}{2}} f_\varphi(p, \lambda)^{\frac{1}{2}}, \end{aligned}$$

providing the convexity of $\ln f_\varphi$ in p . Then for any $(p, \lambda) \in (0, p_\varphi) \times (l_\varphi, L_\varphi)$, we have

$$\begin{aligned} \frac{\partial}{\partial p} \ln F_\varphi(p, \lambda) &= (p+1) \frac{\partial}{\partial p} \ln f_\varphi(p-1, \lambda) + \ln f_\varphi(p-1, \lambda) - p \frac{\partial}{\partial p} \ln f_\varphi(p, \lambda) - \ln f_\varphi(p, \lambda) \\ &\leq p \frac{\partial}{\partial p} \ln f_\varphi(p-1, \lambda) - p \frac{\partial}{\partial p} \ln f_\varphi(p, \lambda) \leq 0 \end{aligned}$$

and deduce the monotonicity of F_φ in p . Since we have

$$\frac{\partial}{\partial \lambda} f_\varphi(0, \lambda) = \exp_\varphi(\lambda), \quad \frac{\partial}{\partial \lambda} f_\varphi(p, \lambda) = p \int_0^{\exp_\varphi(\lambda)} (\lambda - \ln_\varphi(t))^{p-1} \frac{t}{\varphi(t)} dt = p f_\varphi(p-1, \lambda)$$

for $p \in (1, p_\varphi)$, de l'Hôpital's rule with Lemma 3.1(4) yields that

$$\begin{aligned} \lim_{\lambda \downarrow l_\varphi} F_\varphi(1, \lambda) &= \lim_{\lambda \downarrow l_\varphi} 2 \exp_\varphi(\lambda) = 0, & \lim_{\lambda \uparrow L_\varphi} F_\varphi(1, \lambda) &= \lim_{\lambda \uparrow L_\varphi} 2 \exp_\varphi(\lambda) = \infty, \\ \lim_{\lambda \downarrow l_\varphi} F_\varphi(p, \lambda)^{\frac{1}{p}} &= \frac{p^2 - 1}{p^2} \lim_{\lambda \downarrow l_\varphi} F_\varphi(p-1, \lambda)^{\frac{1}{p}}, & \lim_{\lambda \uparrow L_\varphi} F_\varphi(p, \lambda)^{\frac{1}{p}} &= \frac{p^2 - 1}{p^2} \lim_{\lambda \uparrow L_\varphi} F_\varphi(p-1, \lambda)^{\frac{1}{p}} \end{aligned}$$

for any $p \in (1, p_\varphi)$, in particular,

$$\lim_{\lambda \downarrow l_\varphi} F_\varphi(n, \lambda) = \lim_{\lambda \downarrow l_\varphi} F_\varphi(1, \lambda) = 0, \quad \lim_{\lambda \uparrow L_\varphi} F_\varphi(n, \lambda) = \lim_{\lambda \uparrow L_\varphi} F_\varphi(1, \lambda) = \infty$$

for any integer $n \in [1, p_\varphi)$. This with the monotonicity of F_φ in p finishes the proof. \square

Remark 3.3 For the pair of constants (λ, c) obtained for $V = I_d$ in (3.4), we find that $\exp_\varphi(\lambda - c|x - v|_V^2)(\det V)^{-1/2} \mathcal{L}^d$ becomes the probability measure with mean v and covariance matrix V by the change of variables formula for $y = V^{-1/2}(x - v)$.

4 Convexity in Wasserstein geometry

In this section, we investigate the two spaces of φ -exponential distributions with mean and covariance matrix parameters. We mention that $\alpha\varphi(s) = s^q \in \mathcal{O}(2/(d+2))$ with some $\alpha > 0$ if and only if $q \in Q_d := (0, (d+4)/(d+2))$.

Proposition 3.2 and Remark 3.3 yield that for any $\varphi \in \mathcal{O}(2/(d+2))$ with $d \geq 2$, there exist continuous functions λ_φ and c_φ on $S(d, \mathbb{R})_+$ such that each of

$$\begin{aligned} n_\varphi(v, V)(x) &:= \exp_\varphi(\lambda_\varphi(V) - c_\varphi(V)|x - v|_V^2), \\ g_\varphi(v, V)(x) &:= \exp_\varphi(\lambda_\varphi(I_d) - c_\varphi(I_d)|x - v|_V^2)(\det V)^{-\frac{1}{2}} \end{aligned}$$

is the probability density on $(\mathbb{R}^d, \mathcal{L}^d)$ with mean v and covariance matrix V . We define the two spaces of probability measures by

$$\begin{aligned} \mathcal{N}_\varphi &:= \{N_\varphi(v, V) := n_\varphi(v, V)\mathcal{L}^d \mid (v, V) \in \mathbb{R}^d \times S(d, \mathbb{R})_+\}, \\ \mathcal{G}_\varphi &:= \{G_\varphi(v, V) := g_\varphi(v, V)\mathcal{L}^d \mid (v, V) \in \mathbb{R}^d \times S(d, \mathbb{R})_+\}, \end{aligned}$$

which generally differ from each other.

Theorem 4.1 *For any $\varphi, \psi \in \mathcal{O}(2/(d+2))$ with $d \geq 2$, $G_\varphi(0, a^2 I_d) = N_\psi(0, a^2 I_d)$ holds for any $a > 0$ if and only if $\alpha\varphi(s) = \beta\psi(s) = s^q$ with some $\alpha, \beta > 0$ and $q \in Q_d$.*

Proof. We first prove the “if” part. For any $\varphi \in \mathcal{O}(2/(d+2))$ with $d \geq 2$ and $\alpha > 0$, the following relations are verified;

$$\ln_{\alpha\varphi}(t) = \alpha^{-1} \ln_\varphi(t), \quad \exp_{\alpha\varphi}(\tau) = \exp_\varphi(\alpha\tau), \quad f_{\alpha\varphi}(p, \lambda) = \alpha^{-(p+1)} f_\varphi(p, \alpha\lambda).$$

These relations with (3.4) implies that, for any $V \in S(d, \mathbb{R})_+$, we have

$$\begin{aligned} (d\pi)^{-\frac{d}{2}} \frac{\Gamma(d/2)}{\sqrt{\det V}} &= F_{\alpha\varphi}\left(\frac{d}{2}, \lambda_{\alpha\varphi}(V)\right) = F_\varphi\left(\frac{d}{2}, \alpha\lambda_{\alpha\varphi}(V)\right), \\ c_{\alpha\varphi}(V) &= \frac{f_{\alpha\varphi}(d/2, \lambda_{\alpha\varphi}(V))}{df_{\alpha\varphi}((d-2)/2, \lambda_{\alpha\varphi}(V))} = \alpha^{-1} \frac{f_\varphi(d/2, \alpha\lambda_{\alpha\varphi}(V))}{df_\varphi((d-2)/2, \alpha\lambda_{\alpha\varphi}(V))}, \end{aligned}$$

which shows that $\alpha\lambda_{\alpha\varphi}(V) = \lambda_\varphi(V)$ and $\alpha c_{\alpha\varphi}(V) = c_\varphi(V)$, in turn $G_{\alpha\varphi}(v, V) = G_\varphi(v, V)$ and $N_{\alpha\varphi}(v, V) = N_\varphi(v, V)$. Therefore it is enough to prove the case $\varphi(s) = \psi(s) = s^q$ with $q \in Q_d$, which has been already demonstrated in [7, Section 4].

Let us prove the “only if” part. The condition $G_\varphi(0, a^2 I_d) = N_\psi(0, a^2 I_d)$ leads that

$$a^{-d} \exp_\varphi(\lambda - ca^{-2}|x|^2) = \exp_\psi(\lambda(a) - c(a)a^{-2}|x|^2)$$

for any $x \in \mathbb{R}^d$, where we abbreviate $c := c_\varphi(I_d)$, $\lambda := \lambda_\varphi(I_d)$ and $c(a) := c_\psi(a^2 I_d)$, $\lambda(a) := \lambda_\psi(a^2 I_d)$. Evaluating the above equation at $x \in \mathbb{R}^d$ with $|x| = ar$, we deduce that

$$a^{-d} \exp_\varphi(\lambda - cr^2) = \exp_\psi(\lambda(a) - c(a)r^2), \quad b := \sqrt{\frac{\lambda - l_\varphi}{c}} = \sqrt{\frac{\lambda(a) - l_\psi}{c(a)}}.$$

Differentiating this equation at $r \in I := (0, b)$ yields that

$$ca^{-d} \varphi(\exp_\varphi(\lambda - cr^2)) = c(a) \psi(\exp_\psi(\lambda(a) - c(a)r^2)) = c(a) \psi(a^{-d} \exp_\varphi(\lambda - cr^2)),$$

in particular, $c\varphi(\exp_\varphi(\lambda - cr^2)) = c(1)\psi(\exp_\varphi(\lambda - cr^2))$ for the case $a = 1$. This provides

$$\frac{\psi(a^{-d}\exp_\varphi(\lambda - cr^2))}{\psi(\exp_\varphi(\lambda - cr^2))} = \frac{c(1)\psi(a^{-d}\exp_\varphi(\lambda - cr^2))}{c\varphi(\exp_\varphi(\lambda - cr^2))} = \frac{c(1)}{c(a)a^d} \quad (4.1)$$

for any $r \in I$. For any $\xi, \eta > 0$ satisfying $\xi^2, \xi\eta < \exp_\varphi(\lambda)$, since it is possible to choose $s, r \in I$ such that $\xi^2 = \exp_\varphi(\lambda - cr^2)$ and $\xi\eta = \exp_\varphi(\lambda - cs^2)$, the choice $a^{-d} = \eta/\xi$ leads

$$\frac{\psi(\xi\eta)}{\psi(\xi^2)} = \frac{\psi(a^{-d}\exp_\varphi(\lambda - cr^2))}{\psi(\exp_\varphi(\lambda - cr^2))} = \frac{c(1)}{c(a)a^d} = \frac{\psi(\eta^2)}{\psi(\xi\eta)}.$$

Then for any $\alpha \in (0, \min\{\exp_\varphi(\lambda)^{1/2}, \exp_\varphi(\lambda)\})$, the function $\Psi(\xi) := \psi(\xi\alpha^2)/\psi(\alpha^2)$ satisfies that $\Psi(\xi\eta) = \Psi(\xi)\Psi(\eta)$ for any $\xi, \eta \in (0, \alpha^{-1/2}\exp_\varphi(\lambda)^{1/2})$. By induction with the condition $\alpha < \exp_\varphi(\lambda)$, we find that $\Psi(\xi\eta) = \Psi(\xi)\Psi(\eta)$ for any $\xi, \eta > 0$. Hence the continuous function $L(s) := \ln \Psi(e^s)$ on \mathbb{R} satisfies Cauchy's functional equation, namely $L(s+t) = L(s) + L(t)$. This means that $L(s) = L(1)s$, in turn $\alpha\varphi(s) = \beta\psi(s) = s^q$ for some $\alpha, \beta > 0$ and $q \in Q_d$. \square

Remark 4.2 Theorem 4.1 yields that for $\varphi, \psi \in \mathcal{O}(2/(d+2))$ with $d \geq 2$, we have $G_\varphi(0, a^2 I_d) = N_\psi(0, a^2 I_d)$ for any $a > 0$ if and only if \mathcal{G}_φ and \mathcal{N}_ψ are either the space of Gaussian measures, or the space of q -Gaussian measures.

Both the spaces \mathcal{G}_φ and \mathcal{N}_φ are generalizations of the space of Gaussian measures, however they differ from each other in general. On one hand, the space \mathcal{N}_φ has a dually flat structure as seen in [4] (we do not refer to a dually flat structure, see [1]). On the other hand, the space \mathcal{G}_φ is convex in Wasserstein geometry.

Proposition 4.3 *For any $\varphi \in \mathcal{O}(2/(d+2))$ with $d \geq 2$, the space \mathcal{G}_φ is convex and isometric to the space \mathcal{G} of Gaussian measures in the sense of Wasserstein geometry.*

Proof. For $(v, V), (u, U) \in \mathbb{R}^d \times \mathcal{S}(d, \mathbb{R})_+$, we define the symmetric positive definite matrix W and the associated convex function ϕ_W by

$$W := U^{\frac{1}{2}}(U^{\frac{1}{2}}VU^{\frac{1}{2}})^{-\frac{1}{2}}U^{\frac{1}{2}}, \quad \phi_W(x) := \frac{1}{2}\langle x - v, W(x - v) \rangle + \langle x, u \rangle.$$

Then ϕ_W is convex and satisfies $[\nabla\phi_W]_\# G_\varphi(v, V) = G_\varphi(u, U)$ since for $G_\varphi(v, V)$ -almost everywhere we have $g_\varphi(v, V) = g_\varphi(u, U)(\nabla\phi_W)(\det \text{Hess } \phi_W)$. Theorem 2.2 yields that $[\text{id} \times \nabla\phi_W]_\# G_\varphi(v, V)$ is an optimal coupling of $G_\varphi(v, V)$ and $G_\varphi(u, U)$, consequently

$$\begin{aligned} W_2(G_\varphi(v, V), G_\varphi(u, U))^2 &= |v - u|^2 + \text{tr } V + \text{tr } U - 2 \text{tr} \left(U^{\frac{1}{2}} V U^{\frac{1}{2}} \right)^{\frac{1}{2}} \\ &= W_2(G(v, V), G(u, U))^2, \end{aligned}$$

where $G(v, V)$ stands for the Gaussian measure with mean v and covariance matrix V . (Note that the optimality of the function ϕ_W in the case of the Gaussian measures has been known, for instance, see [6] and references therein.) Hence the map $G(v, V) \mapsto G_\varphi(v, V)$ is an isometry from \mathcal{G} to \mathcal{G}_φ in Wasserstein geometry. We also deduce from Theorem 2.2

that, for the time-dependent vector $\{w_t\}_{t \in [0,1]}$ and the time-dependent matrix $\{W_t\}_{t \in [0,1]}$ defined by

$$w_t := (1-t)v + tu, \quad W_t := [(1-t)I_d + tW]V[(1-t)I_d + tW],$$

$\{G_\varphi(w_t, W_t)\}_{t \in [0,1]}$ is a unique Wasserstein geodesic from $G_\varphi(v, V)$ to $G_\varphi(u, U)$. Hence the space \mathcal{G}_φ is convex in Wasserstein geometry. \square

Let us now consider the completions of \mathcal{N}_φ as a metric space with respect to W_2 , denoted by $\overline{\mathcal{N}_\varphi}$, and prove that $\overline{\mathcal{N}_\varphi}$ is generally not convex in Wasserstein geometry. By Proposition 4.3, the completion of \mathcal{G}_φ , denoted by $\overline{\mathcal{G}_\varphi}$, is isometric to the completion of \mathcal{G} , which is homeomorphic to $\mathbb{R}^d \times S(d, \mathbb{R})_{\geq 0}$, where $S(d, \mathbb{R})_{\geq 0}$ is the set of all symmetric non-negative definite matrices of size d (see [6, Section 4]). We similarly show that $\overline{\mathcal{N}_\varphi}$ is homeomorphic to $\mathbb{R}^d \times S(d, \mathbb{R})_{\geq 0}$, whose proof is simple but tedious and we omit it. The key of the proof is the fact that the weak convergence, which is equivalent to the pointwise convergence of characteristic functions, is weaker than the convergence in Wasserstein geometry (see [8, Theorem 7.12]). Hence elements of $\overline{\mathcal{G}_\varphi}$ and $\overline{\mathcal{N}_\varphi}$ are naturally denoted by $G_\varphi(v, V)$ and $N_\varphi(v, V)$ with some $(v, V) \in \mathbb{R}^d \times S(d, \mathbb{R})_{\geq 0}$, respectively. Note that any elements in $\overline{\mathcal{G}_\varphi} \setminus \mathcal{G}_\varphi$ and $\overline{\mathcal{N}_\varphi} \setminus \mathcal{N}_\varphi$ are singular with respect to the Lebesgue measure and $G_\varphi(v, 0) = N_\varphi(v, 0) = \delta_v$, which is the Dirac measure centered at v .

Proposition 4.4 *The space $\overline{\mathcal{N}_\varphi}$ is convex in Wasserstein geometry if and only if $\alpha\varphi(s) = s^q$ with some $\alpha > 0$ and $q \in Q_d$.*

Proof. Since the “if” part is trivially true, we assume the convexity of $\overline{\mathcal{N}_\varphi}$. Then a unique geodesic σ from $N_\varphi(0, 0) = G_\varphi(0, 0)$ to $N_\varphi(0, I_d) = G_\varphi(0, I_d)$ is contained in $\overline{\mathcal{N}_\varphi}$ and $\overline{\mathcal{G}_\varphi}$, which is written as $\sigma(t) = G_\varphi(0, t^2 I_d)$ for $t \in [0, 1]$ and can be extended to $[0, \infty)$. The non-branching property in Wasserstein geometry ensures that $G_\varphi(0, t^2 I_d) \in \mathcal{G}_\varphi \cap \mathcal{N}_\varphi$ for any $t > 0$ and by Theorem 4.1, $\alpha\varphi(s) = s^q$ holds for some $\alpha > 0$ and $q \in Q_d$. \square

5 Stability under an evolution equation

We finally discuss when the space of φ -exponential distributions is stable under the evolution equation of the form

$$\frac{\partial}{\partial t} \rho = \operatorname{div} (\rho \nabla (\ln_\varphi(\rho) + \Psi_\varphi)) = \operatorname{div} \left(\frac{\rho \nabla \rho}{\varphi(\rho)} + \rho \nabla \Psi_\varphi \right), \quad \Psi_\varphi(x) = c_\varphi(I_d) |x|^2. \quad (5.1)$$

We mention that this evolution equation recovers the Fokker–Planck equation (resp. the nonlinear evolution equation of porous medium type) when $\varphi(s) = s$ (resp. $\varphi(s) = s^q$ with some $q \in Q_d$ except for $q = 1$).

A set of functions is said to be *stable* under an evolution equation if any solution to the evolution equation with initial data being an element of the set stays inside the set. A set of absolutely continuous measures with respect to the Lebesgue measure is said to be *stable* under an evolution equation if the set generated by densities of elements in the given set is stable.

Theorem 5.1 Assume $\varphi(0) = 0$. Then the space \mathcal{N}_φ is stable under (5.1) if and only if $\alpha\varphi(s) = s^q$ with some $\alpha > 0$ and $q \in Q_d$.

Proof. Since the “if” part has been already shown in [5, Proposition 5], it is enough to prove the “only if” part, which is done in a similar way to the proof of [5].

Assume the stability of \mathcal{N}_φ . Then the solution to (5.1) with initial data $n_\varphi(0, V) \in \mathcal{N}_\varphi$ is written as $n_\varphi(v_t, V_t)$ for some time-dependent vector $v_t \in \mathbb{R}^d$ and some time-dependent matrix $V_t \in \text{S}(d, \mathbb{R})_+$. The following calculation can be justified;

$$\frac{d}{dt}v_t = \int_{\mathbb{R}^d} x \frac{\partial}{\partial t} n_\varphi(v_t, V_t) d\mathcal{L}^d = - \int_{\mathbb{R}^d} n_\varphi(v_t, V_t) \nabla (\ln_\varphi(n_\varphi(v_t, V_t)) + \Psi_\varphi) d\mathcal{L}^d = -2c_\varphi(I_d)v_t,$$

which implies that v_t is identically the zero vector. We similarly compute

$$\frac{d}{dt}V_t = 4A_tV_t, \quad \frac{d}{dt}V_t^{-1} = -4V_t^{-1}A_t, \quad A_t := c(V_t)V_t^{-1} - c_\varphi(I_d)I_d. \quad (5.2)$$

In what follows, we abbreviate $n_\varphi(0, X)$ as $n_\varphi(X)$. Let ∂_{ij} be the tangent vector on $\mathcal{N}_\varphi(0) := \{n_\varphi(X) \mid X \in \text{S}(d, \mathbb{R})_+\}$ associated to a global chart $(\text{S}(d, \mathbb{R})_+, \Xi) = (\Xi_{ij})_{1 \leq i \leq j \leq d}$ given by $\Xi(X) := n_\varphi(X^{-1})$. We define the map σ on $\text{S}(d, \mathbb{R})_+$ by $\sigma(X) = X^{-1}$ and the symmetric matrices Λ, C of size d by

$$\Lambda := \begin{cases} \partial_{ij}(\lambda_\varphi \circ \sigma)|_{V_t^{-1}} & \text{if } i = j, \\ \frac{1}{2}\partial_{ij}(\lambda_\varphi \circ \sigma)|_{V_t^{-1}} & \text{if } i \neq j, \end{cases} \quad C := \begin{cases} \partial_{ij}(c_\varphi \circ \sigma)|_{V_t^{-1}} & \text{if } i = j, \\ \frac{1}{2}\partial_{ij}(c_\varphi \circ \sigma)|_{V_t^{-1}} & \text{if } i \neq j. \end{cases}$$

Given any $x \in \mathbb{R}^d$, we set the function e_x on $\mathcal{N}_\varphi(0)$ as $e_x(n_\varphi(X)) = n_\varphi(X)(x)$, which is C^1 owing to $\varphi(0) = 0$. Then for the solution $n_\varphi(V_t)$ to (5.1), we have

$$\begin{aligned} \frac{d}{dt}e_x(n_\varphi(V_t)) &= \frac{\partial}{\partial t}n_\varphi(V_t)(x) = \text{div}(n_\varphi(V_t)\nabla(\ln_\varphi(n_\varphi(V_t)) + \Psi_\varphi))(x) \\ &= 4\varphi(n_\varphi(V_t)(x))c(V_t)\langle x, V_t^{-1}A_tx \rangle - 2n_\varphi(V_t)(x)\text{tr}(A_t). \end{aligned}$$

On the other hand, by using (5.2), a direct computation provides

$$\begin{aligned} \frac{d}{dt}e_x(n_\varphi(V_t)) &= \sum_{1 \leq i \leq j \leq d} \partial_{ij}(e_x \circ \Xi_{ij})|_{V_t^{-1}} \frac{d\langle e_i, V_t^{-1}e_j \rangle}{dt} \\ &= -4\varphi(n_\varphi(V_t)(x)) [\text{tr}(V_t^{-1}A_t\Lambda) - \text{tr}(V_t^{-1}A_tC)|x|_{V_t}^2 - c_\varphi(V_t)\langle x, V_t^{-1}A_tx \rangle]. \end{aligned}$$

We therefore obtain

$$n_\varphi(V_t)(x)\text{tr}(A_t) = 2\varphi(n_\varphi(V_t)(x)) [\text{tr}(V_t^{-1}A_t\Lambda) - \text{tr}(V_t^{-1}A_tC)|x|_{V_t}^2] \quad (5.3)$$

for any $x \in \mathbb{R}^d$. Set

$$\beta := \frac{2\text{tr}(V_t^{-1}A_t\Lambda)}{\text{tr}(A_t)}, \quad \gamma := \frac{2\text{tr}(V_t^{-1}A_tC)}{c_\varphi(V_t)\text{tr}(A_t)}, \quad J_{V_t} := (0, \lambda_\varphi(V_t) - l_\varphi).$$

At $x \in \mathbb{R}^d$ with $c_\varphi(V_t)|x|_{V_t}^2 = r \in J_{V_t}$, the equation (5.3) is deformed as

$$\exp_\varphi(\lambda_\varphi(V_t) - r) = \varphi(\exp_\varphi(\lambda_\varphi(V_t) - r)) (\beta + \gamma r) = - \left(\frac{d}{dr} \exp_\varphi(\lambda_\varphi(V_t) - r) \right) (\beta + \gamma r)$$

and integrating it with respect to $r \in J_{V_t}$ yields

$$\exp_\varphi(\lambda_\varphi(V_t) - r) = \exp_\varphi(\lambda_\varphi(V_t)) \left(1 + \frac{\gamma r}{\beta} \right)^{-\frac{1}{\gamma}}.$$

Taking \ln_φ and differentiating it in r , we have

$$\beta \exp_\varphi(\lambda_\varphi(V_t))^\gamma \varphi(s) = s^{1+\gamma},$$

where we set $s = s(r) := \exp_\varphi(\lambda_\varphi(V_t) - r)$. This implies that $\alpha\varphi(s) = s^q$ holds on $(0, \exp_\varphi(\lambda(V_t)))$ with $\alpha = \alpha(V_t) := \beta \exp_\varphi(\lambda_\varphi(V_t))^\gamma$ and $q = q(V_t) := 1 + \gamma$. We mention that α and q do not depend on V_t since we have $\varphi(s) = \alpha(V_t)^{-1} s^{q(V_t)} = \alpha(U_t)^{-1} s^{q(U_t)}$ on $(0, \min\{\exp_\varphi(\lambda(V_t)), \exp_\varphi(\lambda(U_t))\}) \neq \emptyset$ for any $V, U \in \mathcal{S}(d, \mathbb{R})_+$. Letting $V \rightarrow 0$ and $t \rightarrow 0$, we have $\lambda_\varphi(V_t) \rightarrow L_\varphi$ then $\alpha\varphi(s) = s^q$ for any $s > 0$. \square

The same result holds for \mathcal{G}_φ . The proof is exactly similar but tedious and we omit it.

Corollary 5.2 *Let φ be C^1 and $\varphi(0) = 0$. Then the space \mathcal{G}_φ is stable under (5.1) if and only if $\alpha\varphi(s) = s^q$ with some $\alpha > 0$ and $q \in Q_d$.*

References

- [1] S. Amari, *Differential-geometrical methods in statistics*, Lecture Notes in Statistics, vol. 28, Springer-Verlag, New York, 1985.
- [2] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), no. 4, 375–417.
- [3] M. Knott and C. S. Smith, *Note on the optimal transportation of distributions*, J. Optim. Theory Appl. **52** (1987), no. 2, 323–329.
- [4] J. Naudts, *Estimators, escort probabilities, and ϕ -exponential families in statistical physics*, J. Inequal. Pure Appl. Math. **5** (2004), no. 4, 102, 15pp.
- [5] A. Ohara and T. Wada, *Information geometry of q -Gaussian densities and behaviors of solutions to related diffusion equations*, J. Phys. A **43** (2010), no. 3, 035002, 18pp.
- [6] A. Takatsu, *Wasserstein geometry of Gaussian measures*, to appear in Osaka J. Math. **48** (2011), no. 4.
- [7] A. Takatsu, *Wasserstein geometry of porous medium equation*, preprint (2010). Available at <http://sites.google.com/site/asukatakatsu/home/paper>.

- [8] C. Villani, *Topics in optimal transportation*, American Mathematical Society, Providence, RI, 2003.
- [9] C. Villani, *Optimal transport, old and new*, Springer-Verlag, Berlin, 2009.